

DIALOG(R) File 351:Derwent WPI  
(c) 2001 Derwent Info Ltd. All rts. reserv.'

012229747

WPI Acc No: 1999-035854/199904

XRPX Acc No: N99-026892

Pattern recognition method for speech or written data - using hidden Markov model as basis for computation process and transforming characteristic vectors into symbol vectors using vector quantisation

Patent Assignee: DAIMLER-BENZ AG (DAIM )

Inventor: FRANKE J; GLOGER J; KALTENMEIER A; MANDLER E

Number of Countries: 001 Number of Patents: 001

Patent Family:

| Patent No   | Kind | Date     | Applicat No | Kind | Date     | Week     |
|-------------|------|----------|-------------|------|----------|----------|
| DE 19723294 | A1   | 19981210 | DE 1023294  | A    | 19970604 | 199904 B |

Priority Applications (No Type Date): DE 1023294 A 19970604

Patent Details:

| Patent No   | Kind | Lan | Pg | Main IPC    | Filing Notes |
|-------------|------|-----|----|-------------|--------------|
| DE 19723294 | A1   |     | 4  | G06K-009/66 |              |

Abstract (Basic): DE 19723294 A

The pattern recognition method is based upon hidden Markov models and is used for the identification of spoken and written information. Characteristic vectors are transformed into symbol vectors using vector quantisation. The quantisation is made using polynomial classifiers in a mathematical process that does not involve iteration.

USE - For speech or letter pattern recognition.

ADVANTAGE - Provides mathematical efficiency with reduced error rate.

Dwg.0/0

Title Terms: PATTERN; RECOGNISE; METHOD; SPEECH; WRITING; DATA; HIDE; MARKOV; MODEL; BASIS; COMPUTATION; PROCESS; TRANSFORM; CHARACTERISTIC; VECTOR; SYMBOL; VECTOR; VECTOR; QUANTUM

Derwent Class: P86; T04; W04

International Patent Class (Main): G06K-009/66

International Patent Class (Additional): G10L-005/06

File Segment: EPI; EngPI

Manual Codes (EPI/S-X): T04-D04; W04-V01



①⑨ **BUNDESREPUBLIK  
DEUTSCHLAND**



**DEUTSCHES  
PATENT- UND  
MARKENAMT**

⑫ **Offenlegungsschrift**  
⑩ **DE 197 23 294 A 1**

⑤① Int. Cl.<sup>6</sup>:  
**G 06 K 9/66**  
G 10 L 5/06

②① Aktenzeichen: 197 23 294.9  
②② Anmeldetag: 4. 6. 97  
④③ Offenlegungstag: 10. 12. 98

**DE 197 23 294 A 1**

⑦① Anmelder:  
Daimler-Benz Aktiengesellschaft, 70567 Stuttgart,  
DE

⑦② Erfinder:  
Franke, Jürgen, 89075 Ulm, DE; Gloger, Joachim,  
89346 Bibertal, DE; Mandler, Eberhard, 89075 Ulm,  
DE; Kaltenmeier, Alfred, Dr., 89075 Ulm, DE

⑤⑥ Für die Beurteilung der Patentfähigkeit in Betracht  
zu ziehende Druckschriften:

DE 42 40 978 A1  
US 55 02 774  
US 54 65 321  
US 52 74 739  
EP 06 27 726 A1

KALTENMEIER, Alfred u.a.: Hidden Markov Models -  
A Unified Approach to Recognition of Spoken and  
Written Language, Mustererkennung 1993,  
15.DAGM-  
Symposium, Lübeck, 27.-29. Sept. 1993, Springer-  
Verlag, Berlin u.a., S.191-198;

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen**

⑤④ **Mustererkennungsverfahren**

⑤⑦ Für ein Mustererkennungsverfahren auf der Basis von  
Hidden-Markov-Modellen wird vorgeschlagen, bei der  
Transformation der Merkmalsvektoren in Symbolvektoren  
anstelle der gebräuchlichen teil-kontinuierlichen Vektor-  
quantisierer mit angenommener statistischer Verteilungs-  
funktion einen Polynom-Klassifikator als Vektor-  
quantisierer einzusetzen, der die tatsächliche Verteilung  
im Merkmalsvektor-Raum besser annähert und zu einer  
deutlichen Reduktion der Fehlerrate des Erkennungssy-  
stems führt.

**DE 197 23 294 A 1**

## Beschreibung

Die Erfindung betrifft ein Mustererkennungsverfahren nach dem Oberbegriff des Patentanspruchs 1.

Für Mustererkennungsverfahren, insbesondere zur Erkennung von Sprache oder von Schrift sind seit längerem bevorzugt Erkennungssysteme auf der Basis von Hidden-Markov-Modellen im Einsatz, siehe beispielsweise [KCGM 93]. Bei derartigen Erkennungs-Systemen werden von Prüfobjekten mehrdimensionale Merkmalsvektoren gewonnen und auf Symbole eines mehrdimensionalen Symbolvektorraumes abgebildet. Die Liste dieser Symbole, die verschiedenen Zuständen des zugrunde gelegten HMM-Modells zugeordnet sind, wird als Codebook bezeichnet. Diese Symbole können wiederum als Koeffizienten eines mehrdimensionalen Symbolvektors aufgefaßt werden.

Für die Transformation der Merkmalsvektoren in den Symbolvektorraum, die als Vektorquantisierung bezeichnet wird, wird prinzipiell zwischen kontinuierlichen, teilkontinuierlichen und diskreten Modellen unterschieden. Die kontinuierlichen Modelle kommen wegen des damit verbundenen Verarbeitungsaufwands in den meisten Fällen nicht in Betracht.

Bei dem diskreten Modell findet eine sogenannte harte Entscheidung bei der Zuordnung eines Merkmalsvektors zu einem Symbol statt. Dem gegenüber wird bei den teilkontinuierlichen Modellen ein Merkmalsvektor mehreren Symbolen mit im Regelfall unterschiedlicher Gewichtung zugeordnet. Für die Bestimmung dieser Gewichte wird beispielsweise für die Cluster im zustandsabhängigen Merkmalsvektor-Raum eine statistische Verteilung, insbesondere eine Gauss- oder eine Gamma-Verteilung, angenommen. Die teilkontinuierlichen Modelle zeigen im allgemeinen eine wesentlich bessere Erkennungsrate als die diskreten Modelle.

Der vorliegenden Erfindung liegt die Aufgabe zugrunde, die Erkennungsleistung eines Mustererkennungsverfahrens dieser Art weiter zu verbessern, ohne den Aufwand bei der Erkennung nennenswert zu steigern.

Die Erfindung ist im Patentanspruch 1 beschrieben. Die Unteransprüche enthalten vorteilhafte Ausgestaltungen und Weiterbildungen der Erfindung.

Die Erfindung geht von der Erkenntnis aus, daß die bei üblichen teilkontinuierlichen Modellen angenommene statistische Verteilung die tatsächliche Verteilung der Merkmalsvektoren im zustandsabhängigen Vektorraum häufig nur sehr unbefriedigend beschreibt. Demgegenüber wird bei der vorliegenden Erfindung keinerlei Annahme über eine bestimmte Verteilungsfunktion gemacht, sondern es werden aus repräsentativen Trainingsdaten die Koeffizienten eines Polynom-Klassifikators bestimmt, welcher bei der Vektorquantisierung eine wesentlich genauere Approximation an die tatsächlich vorliegende Verteilung im Vektorraum darstellt.

In einer vorteilhaften Ausführungsform wird der Polynom-Klassifikator mit Momenten-Matrizen trainiert, die auch bei einem Training der HMM-Modelle anfallen. Die Vektorquantisierung erfolgt vorzugsweise mehrstufig verzweigt in der Art, daß mehrere Polynom-Teil-Klassifikatoren vorgesehen sind, welche die Knoten der Verzweigung darstellen und in einer baumartigen Verzweigung angeordnet sind, was gegenüber einem einzigen Klassifikator den Vorteil hat, daß eine Beschränkung auf nur die besten Pfade des Baumes vorgenommen werden kann und dadurch wesentlich weniger Polynome evaluiert werden müssen. Vorzugsweise sind diese Teil-Klassifikatoren als binäre Polynom-Klassifikatoren ausgeführt. Vorteilhafterweise wird in an sich bekannter Art vor der Durchführung der Vektor-

quantisierung die Dimension der Merkmalsvektoren mittels einer linearen Diskriminanz-Analyse reduziert.

Die Erfindung ist nachfolgend noch anhand eines Erkennungssystems zur Erkennung gebundener Schrift veranschaulicht. Ein solches Erkennungssystem ist beispielsweise in [KCGM 93] beschrieben und für Einzelheiten eines solchen beispielhaften Erkennungssystems wird auf diese Literaturstelle verwiesen.

Ausgehend von einem binären Pixel-Bild einer handgeschriebenen Buchstaben- oder Zahlenfolge besteht eine wesentliche Aufgabe beim Einsatz eines HMM-Erkennters darin, aus dem Binär-Bild eine Folge von Vektoren als Eingangsgrößen des HMM-Erkennters zu erzeugen. Eine solche Vorverarbeitung kann beispielsweise den Übergang von der Pixel-Darstellung zu einer Kontur-Beschreibung umfassen, an den sich dann Normierungsschritte, wie Schreiblinienschätzung, Aufrichtung schräger Schrift, Drehung des Schriftbilds und Übergang zu einer Skelett-Darstellung anschließen können. Solche Vorverarbeitungsmaßnahmen sind in verschiedener Weise bekannt und gebräuchlich. Eine detailliertere Darstellung der Vorverarbeitungs-Maßnahmen eines solchen Erkennungssystems ist beispielsweise in [CGM 93] beschrieben. Die Einstellung eines HMM-Erkennters ist komplex und, da in verschiedenen Ausführungen aus dem Stand der Technik hinreichend bekannt, an dieser Stelle nicht im Detail beschrieben. Zur Erzeugung von Merkmalsvektoren aus dem vorverarbeiteten Bild wird beispielsweise ein schmaler Fensterausschnitt in angenommener Schreibrichtung schrittweise über das Schriftbild bewegt und bei jedem Schritt werden aus dem Fensterinhalt Merkmale extrahiert, die in ihrer Gesamtheit einen Merkmalsvektor bilden. Diese Vorgehensweise wird sowohl in einer Trainingsphase auf Schriftbilder bekannten Inhaltes als auch in einer nachfolgenden Erkennungsphase auf Schriftbilder unbekannten Inhaltes angewandt.

Bei ausreichend großer Anzahl von Trainingsmustern zeigen die aus diesen Mustern gewonnenen Merkmalsvektoren Anhäufungen, sogenannte Cluster, im mehrdimensionalen Raum der Merkmalsvektoren. Diese Cluster können auf an sich bekannte Weise, beispielsweise nach dem bekannten und überwiegend angewandten LBG-Algorithmus analysiert werden.

Gemäß einem aus dem Stand der Technik bekannten vorteilhaften Vorgehen werden bei Bewegung des Fensterausschnitts nicht nur die statischen Merkmale einer Fensterposition, sondern auch die Differenzen der Merkmale zu benachbarten Fensterpositionen bestimmt und weiterverarbeitet, was zu einem entsprechend höherdimensionalen Merkmalsvektor führt. Die nachfolgenden Ausführungen gelten, ohne daß es jeweils eines gesonderten expliziten Hinweises bedarf, auch für solche und ähnliche Varianten der Merkmalsgewinnung.

In einer ersten Trainingsphase des HMM-Erkennters werden auf der Basis der zuvor erfolgten zustandsunabhängigen Vektorquantisierungstrainings (ein oder mehrere sogenannte Codebooks) und unter Zugrundelegens vorgebarbarer Strukturen der HMM-Modelle für die möglichen Objektklassen mögliche Pfade durch die Strukturen und Zuordnungen einzelner Merkmalsvektoren zu Zuständen der HMM-Modelle, z. B. unter Anwendung des sogenannten Forward-Backward-Verfahrens, bestimmt und dabei die Koeffizienten der HMM-Modelle ermittelt. In diesem Rahmen wurden auch für die verschiedenen Zustände, die die Symbole eines einzelnen neuen zustandsabhängigen Codebooks bilden, Momentenmatrizen aus den zuzuordnenden Merkmalsvektoren bestimmt. Die Momentenmatrizen werden zur Schätzung einer linearen Diskriminanz-Transformations-Matrix herangezogen, die wiederum zur Reduzierung der Dimension der

Merkmalen dient. Diese Verfahrensschritte sind an sich bekannt und gebräuchlich und im Detail in (Fuk 90) beschrieben.

Aus den Momentenmatrizen können bei Einsatz eines Vektorquantisierers auf der Basis angenommener statistischer Verteilungen, z. B. Gauß-Verteilungen, die Parameter der Verteilung wie z. B. Schwerpunkt und Kovarianz ermittelt werden. Solche angenommenen statistischen Verteilungen nähern jedoch die tatsächlichen Verteilungen häufig nur unzureichend an.

Gemäß der vorliegenden Erfindung wird mit dem Einsatz eines Polynom-Klassifikators ein gänzlich anderer Weg beschritten, bei welchem a priori keinerlei Annahmen über eine bestimmte statistische Verteilung gemacht werden. Die Koeffizienten des Polynom-Klassifikators werden auf der Basis der in der Trainingsphase bestimmten Merkmalsvektoren ermittelt.

Vorteilhafterweise werden dabei anstelle der Merkmalsvektoren die bereits zuvor gebildeten Momentenmatrizen herangezogen und so der bei der Koeffizientenermittlung ohnehin erforderliche Schritt der Bildung dieser Matrizen aus den Merkmalsvektoren vermieden. Darüber hinaus kann auf diesem Weg vorteilhafterweise die explizite Kennzeichnung der Klassenzugehörigkeit der einzelnen Merkmalsvektoren umgangen werden, da die Momentenmatrizen ohnehin zu bestimmten HMM-Zuständen gehören und diese im späteren Vektorquantisierer auf Polynomklassifikatorbasis gerade die zu erzeugenden Symbole darstellen.

Aus Gründen des Rechenaufwands ist es auch hier vorteilhaft, mit der gegenüber dem ursprünglichen Merkmalsvektor-Raum reduzierten Dimension zu arbeiten, indem die lineare Diskriminanz-Transformation auf die Merkmalsvektoren bzw. die Momentenmatrizen angewandt wird.

Die Polynomstruktur wird vorzugsweise als vollständige Struktur vorgegeben, kann aber auch zur Reduzierung des Rechenaufwands ausgedünnt werden und bleibt während der Trainingsphase unverändert. Für die Diskriminanzfunktion  $d(v)$  des Polynoms zu den verschiedenen Objektklassen gilt

$$d(v) = A^T x(v)$$

mit  $A$  als Koeffizientenmatrix und  $x(v)$  als Funktion für eine vektorielle Abbildung des ursprünglichen Merkmalsvektors  $v$  auf die Polynomstrukturliste  $x$ .

Für jeden Knoten des Pfad-Verzweigungsbaums (trellis) kann die Wahrscheinlichkeit, daß ein Merkmalsvektor einem bestimmten Zustand eines HMM-Modells zuzuordnen ist, bestimmt werden, was zu neuen Momentenmatrizen  $E\{xx^T\}$  und  $E\{xy^T\}$  mit  $y$  als Symbolklassen-Vektor führt, aus denen die Koeffizientenmatrix nach

$$A = E\{xx^T\}^{-1} E\{xy^T\}$$

bestimmt werden kann. Eine detaillierte Beschreibung dieser und weiterer Aspekte von Polynom-Klassifikatoren gibt [Sch 96].

Die Verwendung eines Polynom-Klassifikators bietet, zumindest bei Verwendung des kleinsten mittleren Fehlerquadrats als Optimierungskriterium, den Vorteil einer mathematisch geschlossenen Lösung ohne Iterationen.

Zur Verringerung des Rechenaufwands wird die Vektorquantisierung vorzugsweise nicht mittels eines einzigen einstufigen Polynom-Klassifikators, sondern mehrstufig mit sukzessiver Verzweigung in aufeinanderfolgenden Stufen unter Einsatz mehrerer Polynom-Teil-Klassifikatoren nach Art einer Baumstruktur durchgeführt. Die Teilklassifikatoren sind insbesondere vorteilhafterweise als binäre Poly-

nom-Klassifikatoren ausgeführt.

Die Einstellung des Polynom-Klassifikators in einer Trainingsphase wird im Regelfall gegenüber einem Vektorquantisierer mit angenommener statistischer Verteilungsfunktion in dieser Trainingsphase mit einem höheren Rechenaufwand verbunden sein. Dieser fällt jedoch nur einmal in der Trainingsphase an. Der Rechenaufwand in der späteren Erkennungsphase ist nicht oder nur unwesentlich höher im Vergleich zur Vektorquantisierung auf der Basis statistischer Verteilungsfunktionen. Bei Durchlaufen des Verzweigungsbaums der Teil-Klassifikatoren kann durch Abbrechen unbedeutender Pfade (pruning) der Verarbeitungsaufwand bei der Vektorquantisierung weiter reduziert werden.

Die verbesserte Beschreibung der tatsächlichen Verteilung im Merkmalsvektor-Raum durch den Polynom-Klassifikator führt nach ersten Untersuchungen gegenüber einer angenommenen Normalverteilung zu einer signifikanten Verringerung der Fehlerrate, bei sonst gleicher Einstellung des Erkennungssystems beispielsweise zu einer Fehlerrate von 5,6% gegenüber 6,3% bei angenommener Normalverteilung.

#### Literaturliste

- [KCGM93] A. Kaltenmeier, T. Caesar, J.M. Gloger, and E. Mandler. Sophisticated topology of hidden markov models for cursive script recognition. In Proceedings of the Second International Conference on Document Analysis and Recognition, S. 139-142, Tsukuba Science City, Japan, 20.-22. Oktober 1993. IEEE Computer Society Press
- [CGM93] T. Caesar, J.M. Gloger, and E. Mandler. Preprocessing and feature extraction for a handwriting recognition system. In Proceedings of the Second International Conference on Document Analysis and Recognition, S. 408-411 Tsukuba Science City, Japan, 20.-22. Oktober 1993, IEEE Computer Society Press.
- [Fuk90] Keinosuke Fukunaga, Introduction to statistical pattern recognition, Academic Press Inc., San Diego, CA, 2. Edition, 1990
- [Sch96] Jürgen Schürmann, Pattern Classification, John Wiley & Sons Inc., New York 1996

#### Patentansprüche

1. Mustererkennungsverfahren auf der Basis von Hidden-Markov-Modellen, bei welchem Merkmalsvektoren mittels einer Vektorquantisierung in Symbol-Vektoren transformiert werden, **dadurch gekennzeichnet**, daß die Vektorquantisierung mittels eines Polynom-Klassifikators ohne Annahme über im zustandsabhängigen Merkmalsraum vorliegende statistische Verteilungen durchgeführt wird.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Vektorquantisierung in mehreren Stufen verzweigt erfolgt wobei einzelne Polynom-Teil-Klassifikatoren die Knoten der Verzweigung repräsentieren.
3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß die Polynom-Teil-Klassifikatoren binäre Polynom-Klassifikatoren sind.
4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß die Dimension der Merkmalsvektoren reduziert und die Vektorquantisierung anhand der dimensionsreduzierten Merkmalsvektoren durchgeführt wird.
5. Verfahren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß die Koeffizienten des Polynom-Klassifikators aus den in einer Trainingsphase ermittelten Merkmalsvektoren oder daraus abgeleiteten

Größen geschätzt werden.

6. Verfahren nach Anspruch 5, dadurch gekennzeichnet, daß in einer Trainingsphase des HMM-Erkenner aus den Merkmalsvektoren Momenten-Matrizen gebildet und diese Momenten-Matrizen auch zur Schätzung der Koeffizienten des Polynom-Klassifikators, bzw. der Polynom-Klassifikatoren eingesetzt werden.

10

15

20

25

30

35

40

45

50

55

60

65